

Article

# Research on Seabed Sediment Classification Based on the MSC-Transformer and Sub-Bottom Profiler

Han Wang <sup>1,†</sup> , Qingjie Zhou <sup>2,†</sup>, Shuo Wei <sup>1</sup> , Xiangyang Xue <sup>1</sup> , Xinghua Zhou <sup>2</sup> and Xiaobo Zhang <sup>1,\*</sup> 

<sup>1</sup> College of Ocean Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China; wanghan@sdust.edu.cn (H.W.); ws13869422596@sdust.edu.cn (S.W.); xxy@sdust.edu.cn (X.X.)

<sup>2</sup> First Institute of Oceanography, Ministry of Natural Resources, Qingdao 266061, China; zhouqj@fio.org.cn (Q.Z.); xhzhou@fio.org.cn (X.Z.)

\* Correspondence: zxb@sdust.edu.cn

† These authors contributed equally to this work.

**Abstract:** This paper proposed an MSC-Transformer model based on the Transformer's neural network, which was applied to seabed sediment classification. The data came from about 2900 km<sup>2</sup> of seabed area on the northern slope of the South China Sea. Using the submarine backscattering intensity and depth data obtained by the sub-bottom profiler, combined with latitude and longitude information, a seabed dataset of the slope area of the South China Sea was constructed. Moreover, using the MSC-Transformer, the accurate identification and judgment of sediment types such as calcareous bio-silt, calcareous bio-clay silt, silty sand, medium sand and gravel sand were realized. Compared with the conventional deep neural network CNN, RNN, etc., the model shows advantages when applied to the sediment dataset of the shallow sea slope region of the South China Sea. This confirms the feasibility and validity of the model and provides a reliable and accurate tool for seabed sediment classification in the field of marine science. The completeness and accuracy of the dataset and the good performance of the model provide a solid foundation for the scientificity and practicability of the study.

**Keywords:** seabed sediment classification; sub-bottom profiler; deep learning; machine learning; transformer model



**Citation:** Wang, H.; Zhou, Q.; Wei, S.; Xue, X.; Zhou, X.; Zhang, X. Research on Seabed Sediment Classification Based on the MSC-Transformer and Sub-Bottom Profiler. *J. Mar. Sci. Eng.* **2023**, *11*, 1074. <https://doi.org/10.3390/jmse11051074>

Academic Editor: Alfredo L. Aretxabaleta

Received: 26 April 2023

Revised: 12 May 2023

Accepted: 14 May 2023

Published: 18 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Seabed sediment type is a key marine environmental parameter and an integral component in ocean exploration [1]. It is of great significance to ascertain the complex types of seabed sediment, especially in the fields of marine technology [2], marine engineering construction, and national defense and military [3]. Seabed sediments usually include gravel, sand, silt, fine sand, mud, reef, etc. [4,5]. The area studied in this paper is the northern slope of the South China Sea, where the seabed terrain fluctuates greatly. Areas with water depths less than 1000 m are mainly affected by the ancient sand belts on the outer edge of the continental shelf, including Silty Sand, Medium Sand, and Gravel Sand, and areas with a water depth greater than 1000 m exhibit semi-deep-sea characteristics, with Calcareous Bio-silt and Calcareous Bio-clay silt being the main ones. Early methods for obtaining types of seabed sediment were in-situation and laboratory measurements [6,7]. Although these methods can directly obtain the types of sediments, the collection ranges are limited due to the influence of sampling methods, so the amount of effective data obtained is small, and the type of seabed sediments in the entire study area cannot be fully displayed [8]. Due to the low efficiency and high cost of in-situation and laboratory measurements, indirect methods are commonly used in actual ocean measurements [9].

The acoustic-based indirect methods have the advantages of high coverage, dense sampling, high efficiency, and low cost, making them widely used in seabed sediment

classification [10–13]. Remote acoustic measurements using acoustic detection equipment such as multi-beam systems and side-scan sonar are commonly used in indirect methods [14–16]. For side-scan sonar, Reut et al. [17] proposed a scheme to extract power spectra from side-scan sonar data and feature values from power spectra for seabed sediment classification in 1985. Duncan et al. [12] established a parameter model for tracking the power spectrum of side-scan sonar and used feature extraction methods to achieve the classification of seabed sediments. Daniel Buscombe et al. [18] used spectral analysis based on short-echo sequence wavelet transform to extract texture features of different attributes from echo maps obtained from side-scan sonar, and used statistical methods to establish the relationship between texture longitudinal scale and sediment types. Afterwards, L. Atallah et al. [19] used data from side-scan sonar to identify the seabed sediments using the most significant wavelet features determined based on backpropagation elimination algorithms. For multi-beam systems, Neil C. Mitchell et al. [10] first combined bathymetry with multi-beam sonar data and used statistical methods to classify seabed sediments. Huseby and Snellen et al. [20] used the backscatter data collected by multi-beam echo detectors to adopt the Bayesian statistical method and K-Means clustering. Then Kazi Ishtiak Ahmed et al. [21] studied principal component analysis and k-means clustering, and proposed an alternative method that utilizes computational and visual data mining to reduce the complexity of the original method. And Tang [22], based on a multi-beam system and sampling data of seabed sediments, established a relationship model between the backscatter intensity of the seabed and the characteristics of sediment types.

From the principles of side-scan sonar and multi-beam system, due to the high frequency and low sound power of the emitted sound waves, the multi-beam system and side-scan can penetrate the sediment layer at a shallow depth, ranging from a few centimeters to more than ten centimeters, making it difficult to obtain sediment information from deeper layers and obtain more feature information to describe the type of sediment [23,24]. Therefore, the authenticity of the obtained data may be problematic.

The sub-bottom profiler is an instrument that analyzes the reflection time, amplitude, frequency, and other information of the received reflected waves based on the differences in the intensity of the reflected sound waves on the seabed, in order to obtain the characteristics and properties of the effective penetration of sound waves into the formation [24–26]. By amplifying the filtered echo signal, a clear cross-section structure of the formation can be depicted using lines composed of different grayscale points. The frequency of sound waves emitted by sub-bottom profilers is usually between a few hundred to tens of thousands of hertz, and they can detect geological structures within a depth of 50 to 200 m [25]. The relationship between the frequency of emitted sound waves and the detection depth can be freely adjusted according to the sediment conditions of different sea areas to obtain the required resolution and accuracy of water depth and reflection intensity [27], such as the northern slope of the South China Sea in this study. Due to the fact that its echo contains more information about shallow sediment, the confidence level set for inferring the type of sediment is higher [28]. LeBlanc et al. [29] proposed a model. By using this model, it can be observed that the pulse signal obtained by the sub-bottom profiler achieves the purpose of bottom material classification through the difference in phase dispersion of different bottom materials. Stevenson et al. [30] extracted the quality factor  $Q$  from the data of the sub-bottom profiler, and compared it with the laboratory results through the acoustic attenuation model Determine the sediment type of the seabed sedimentary layer. Evangelos Alevizos et al. [31] used bathymetric and sub-bottom profiler data and Bayesian methods to classify sediments based on the variability of backscatter at a single incident angle. They also used the maximum likelihood method to classify and principal component analysis multi-angle layers and this method can distinguish acoustic similar categories in different bathymetric environments. Mohamed Saleh et al. [32] improved the model originally used for high-frequency signal classification, taking into account the interaction between low-frequency echo sub-layers, and their seabed sediment classification results were basically consistent with the visual inspection results of their samples. Zheng et al. [3]

used seafloor profile data from a survey line on the northern slope of the South China Sea and used the Biot model for inversion, quantitatively estimating the particle size and sediment types of the underlying materials in the relevant area.

However, there are still some limitations in the bottom material classification methods of sub-bottom profiler, side-scan sonar, and multi-beam system in terms of data processing. Traditional sediment classification methods include statistical methods and empirical models, but they usually have drawbacks such as low efficiency, large limitations, poor transferability, and high data dependence [24,33]. Recently, deep learning methods based on artificial neural networks have been widely applied in fields such as object detection, image processing, and data classification [34]. Due to their strong transfer ability and capability to efficiently process large amounts of data, applying deep learning methods to seabed sediment classification has gradually become a research trend [6,35].

Deep learning is a learning method based on neural networks that extracts and learns the features of input data, undergoes multi-layer transformations, and ultimately outputs the results [36]. The core of deep learning is tensor and auto grid, which is achieved through the forward propagation process for auto grid of tensor, and the back propagation algorithm used to extract data features and update node weights. The commonly used deep neural networks are convolutional neural networks (CNN) and recurrent neural networks (RNN). Yang et al. [37] used a CNN method to realize the classification of seabed sediment, which added bathymetric data to the multi-beam backscatter intensity data for auxiliary classification, improving the classification effect. Tim Berthold et al. [38] proposed an automatic classification method for seabed sediments based on CNN side-scan sonar images. Luo et al. [39] validated that CNN classifiers can be applied to small seabed acoustic image datasets, and shallow CNN outperforms deep CNN in classification accuracy and speed. Qin et al. [40] optimized the small dataset generated by side-scan sonar using deeper CNN to improve classification accuracy in response to the problem of limited datasets.

However, CNN has limited capacity and lacks the ability to handle large amounts of data, and the number of operations required for CNN to calculate the association between two positions increases with distance, which greatly increases the computational complexity [41]. In addition, RNNs have not yet been widely used in the field of seabed sediment classification. Moreover, RNNs cannot perform parallel computation and are prone to gradient vanishing problems, making it unable to handle encoded sequence information [42].

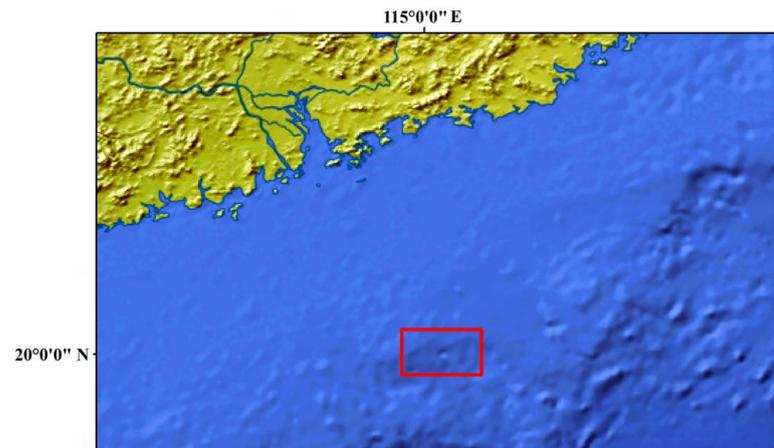
In 2017, Transformer was born, and this S2S model based on attention mechanism caused a sensation in the field of natural language processing (NLP) once it came out [43]. In 2020, DETR [44] demonstrated that Transformer can perform image classification tasks well without CNN. Due to Transformer's outstanding performance in the fields of visual images and semantic analysis, more possibilities are gradually being explored. In December 2020, Amazon AWS [45] proposed TabTransformer, which first combined the characteristics of Transformers such as parallel computing and multi-head attention mechanism, confirming the feasibility of Transformers in processing table text and data. This paper proposes the MSC-Transformer (Transformer of Marine Sediment Classification) based on the TabTransformer method for processing tabular data about reflection intensity data of the northern slope of the South China Sea obtained by the sub-bottom profiler. Compared to other deep learning models, this model has better parallelization ability and larger capacity and performs better when processing larger-scale data [46]. MSC-Transformer can set static hyperparameters, and obtain high-precision classification models without repeated training, and a single NVIDIA GeForce RTX 3090/PCIe/SSE2 GPU can run. It can propose good solutions for the processing of different batches of marine sediment datasets and their different types of division.

In this study, we propose a new acoustic seabed sediment classification method based on MSC-Transformer for sub-bottom profiler reflection intensity, and validate the feasibility and accuracy of this method using reflection intensity data from the northern slope of the South China Sea. The structure of the remaining part of this article is arranged as

follows. Section 2 introduces the construction of the dataset, and Section 3 will elaborate on the proposed model. Section 4 provides an experimental description and result analysis. Finally, a summary and discussion were provided in Section 5.

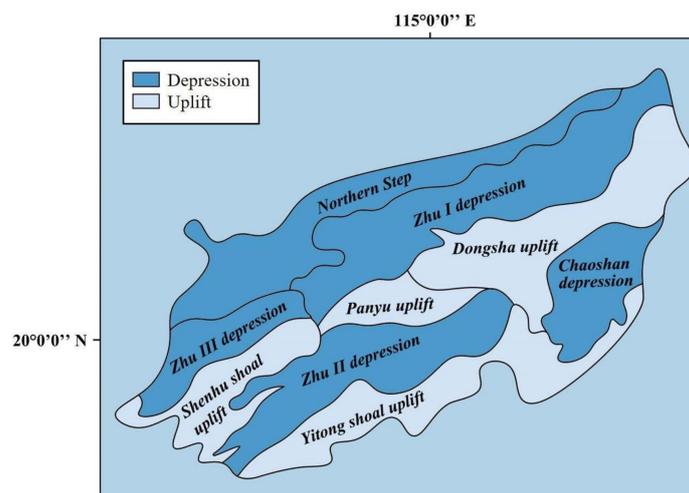
## 2. Research Background

The area of this study is part of the northern slope of the South China Sea, including a seabed ranging from 19°85' N to 20°20' N, and 114° 80' E to 115°60' E, which the total area is approximately 2900 km<sup>2</sup>. The location of the study area is shown in following Figure 1.



**Figure 1.** The location of study area. The red frame represents the study area.

The northern slope of the South China Sea is located at the junction of the Eurasian Plate and the Pacific Plate. Affected by the Tethys-Pacific structural domain, the seismic sections are complex, and many lower lifts and depressions are distributed in the northeast-southwest direction. The area of this study is between Panyu Low Uplift and Zhu II Depression, whose structural geological framework is shown in Figure 2.



**Figure 2.** Structural geological framework of study area.

Because of the complex seismic sections and the long-term internal and external pressure, the topography and geomorphology of the seabed are characterized by great depth, diversity and complexity. The eastern and western ends of the northern slope of the South China Sea are respectively the southeastern end of Taiwan Island and the eastern mouth of the Xisha Trough, spreading from north to east, with the eastern side wider than the western side. And there are numerous submarine geomorphic units, including the continental shelf, the South China Sea Trough, and the Nansha Steps.

### 3. Construction of Seabed Sediment Classification Datasets Based on Sub-Bottom Profiler

#### 3.1. A Method of Constructing Dataset Based on Overlay

To construct a dataset for the northern slope of the South China Sea, this paper adopts the Overlay method. There are  $N$  parallel layers with the same projection coordinate system in the space Cartesian coordinate system, and these layers are represented by  $P_1, P_2 \dots P_{n-1}$ . The area identified on  $P_n$  is  $x_{P_{n-1}min} \leq x_{P_{n-1}} \leq x_{P_{n-1}max}, y_{P_{n-1}min} \leq y_{P_{n-1}} \leq y_{P_{n-1}max}$ . Each layer carries a type of feature, and the point where the feature is located is represented by  $(x_{P_{m1}}, y_{P_{m1}}), (x_{P_{m2}}, y_{P_{m2}}) \dots (x_{P_{mi}}, y_{P_{mi}}), (i \in 1, 2, 3 \dots)$ . Each element contains an attribute, represented by  $V_1, V_2 \dots V_n$ . The processing diagram of the Overlay method is shown in Figure 3.

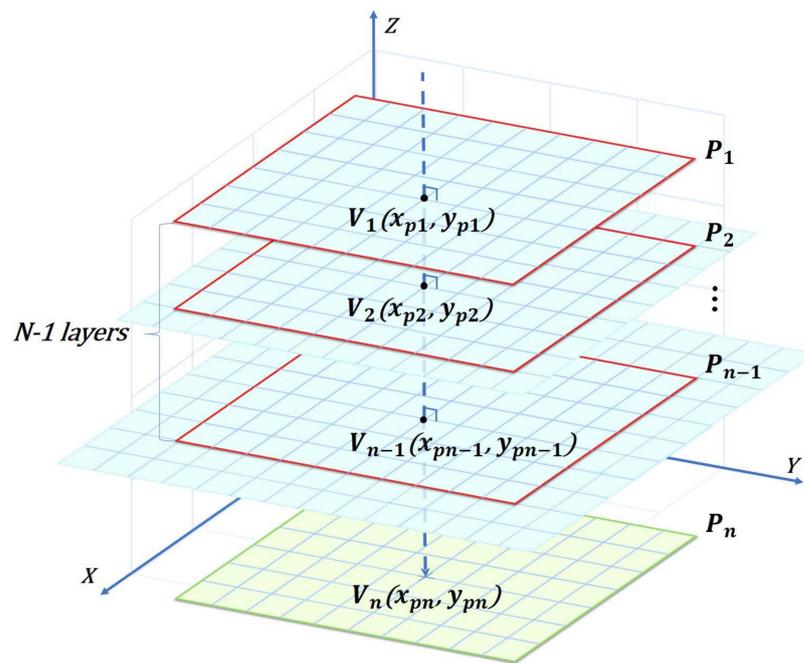


Figure 3. Processing of overlay.

When  $x_{P_i} = x_{P_n}, y_{P_i} = y_{P_n} (i \in 1, 2, 3 \dots)$ , all layers are stacked along the direction of  $z$ -axis and projected onto a new layer  $P_n$ , where values of attribute from different layers are recorded in the form of fields.

#### 3.2. Introduction to Data on Northern Slope of the South China Sea

##### 1. Acquisition of reflection intensity

In this study, a scientific research ship carrying a sub-bottom profiler was used to obtain the reflection intensity of the northern slope of the South China Sea, with a total of 49,998 data points, forming 35 survey lines. Except for one survey line used to measure the submarine valley, the rest of the survey lines are distributed continuously and parallel to each other in a southwest-northeast direction, covering the entire study area. The distribution of survey lines is shown in Figure 4.

The seabed reflection intensity of the study area was obtained by the sub-bottom profiler. Along with the traces of survey lines, the shape of seabed and its times are displayed, and two representative seismic sections are shown in Figure 5; Obtain the positions of these seismic sections in all survey lines as shown in Figure 6.

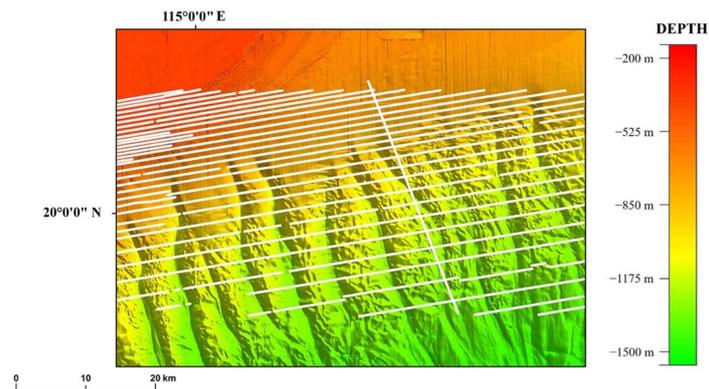


Figure 4. Schematic diagram of the study area and survey lines. White lines represent the survey lines.

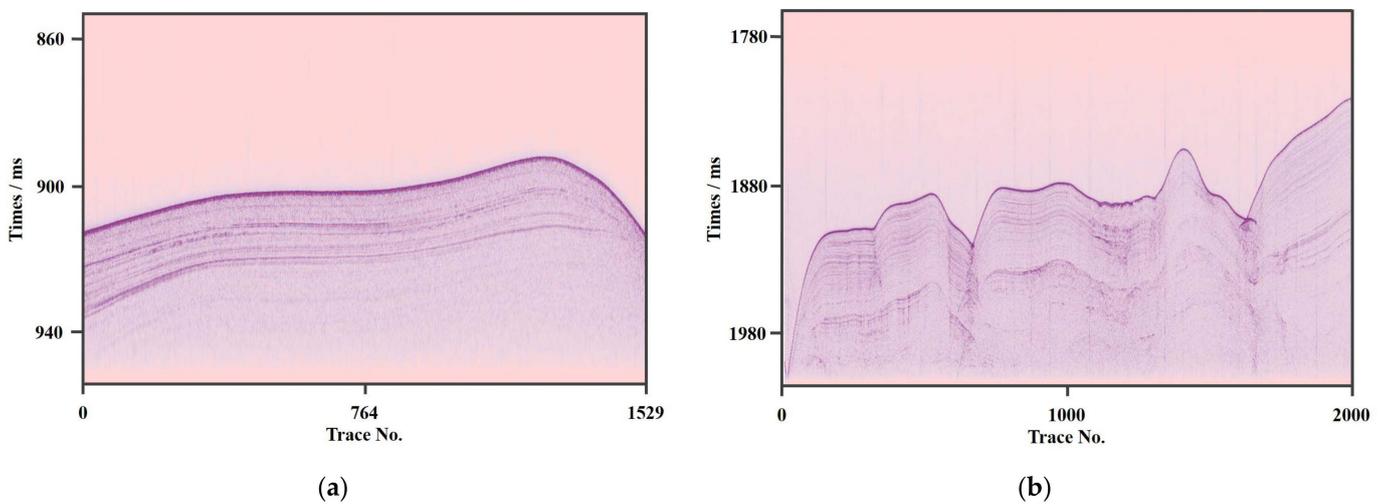


Figure 5. (a) Representative seismic section I; (b) Representative seismic section II.

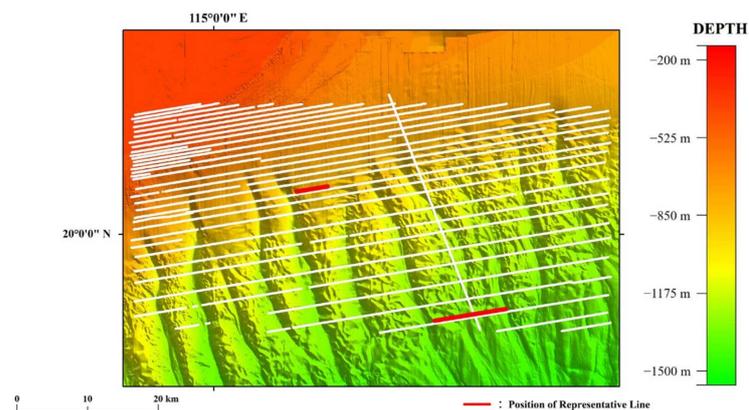


Figure 6. Position of representative lines in all survey lines. White lines represent survey lines, the upper red line is corresponding to Figure 5a and the lower red line is corresponding to Figure 5b.

## 2. Acquisition of water depth

The sub-bottom profiler can obtain the depth of the detected sea area by calculating the time difference between the transmission and reception of sound waves. The topographic map of the study area is shown in Figure 7.

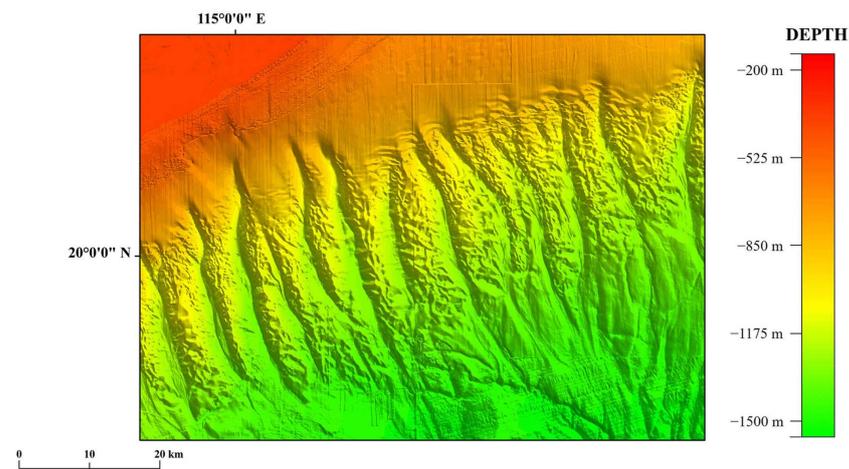


Figure 7. Topographic map of the study area.

### 3. Acquisition of types of seabed sediment

The types of seabed sediment are the label in the dataset, and based on previous research using in-situ measurements, the obtained types of seabed sediment in the study area will be applied as true values in the supervised classification of MSC-Transformer. The map of seabed sediment distribution in the study area is shown in Figure 8.

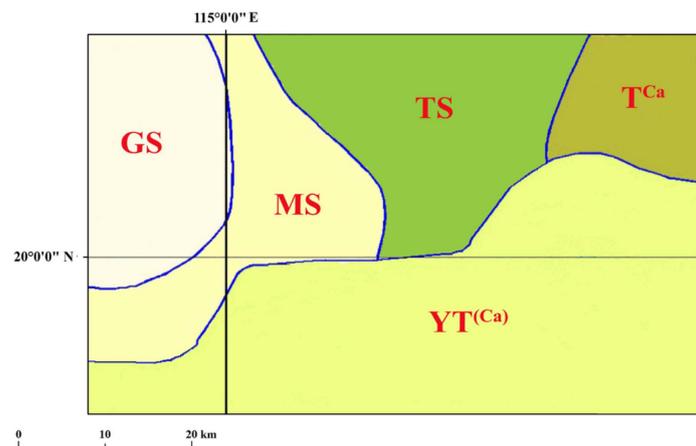
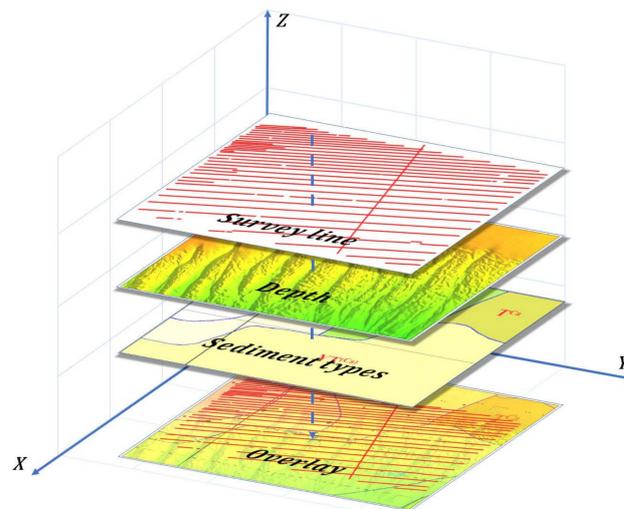


Figure 8. Map of seabed sediment distribution.  $T^{Ca}$  represents Calcareous Bio-silt,  $YT^{(Ca)}$  represents Calcareous Bio-clay silt, TS represents Silty Sand, MS represents Medium Sand, GS represents Gravel Sand.

#### 3.3. Construction of Dataset on Northern Slope of the South China Sea Based on Overlay

Next, the method described in Section 3.1 will be used to construct the South China Sea North Slope data introduced in Section 3.2 to form the required dataset for the study. In this study, WGS-84 is used as the plane coordinate system, Mercator projection is used as the projection coordinate system, the 1985 China National Height Datum is used as the elevation datum and the theoretical depth datum is used as the depth datum. The layers in this study include survey lines, depth, and sediment types, respectively containing elements such as reflection intensity, depth, and sediment type, with attribute values corresponding to the elements. The operation diagram of the Overlay is shown in Figure 9.



**Figure 9.** Operation of overlay. The red lines represent the survey lines.

After removing invalid values and normalization processing, this dataset involves a total of 5 types of seabed sediment, and the label of calcareous bio-silt is “0”, calcareous bio-clay silt is “1”, silty sand is “2”, medium sand is “3” and gravel sand is “4”. The information of seabed sediment data is shown in Table 1, while the format of our dataset is shown in Table 2.

**Table 1.** The information of seabed sediment data.

Types	Abbreviation	Quantity	Gray-Values	Labels
Calcareous Bio-silt	T <sup>Ca</sup>	3915	165	0
Calcareous Bio-clay silt	YT <sup>(Ca)</sup>	7698	197	1
Silty Sand	TS	13,055	242	2
Medium Sand	MS	11,270	243	3
Gravel Sand	GS	14,060	251	4

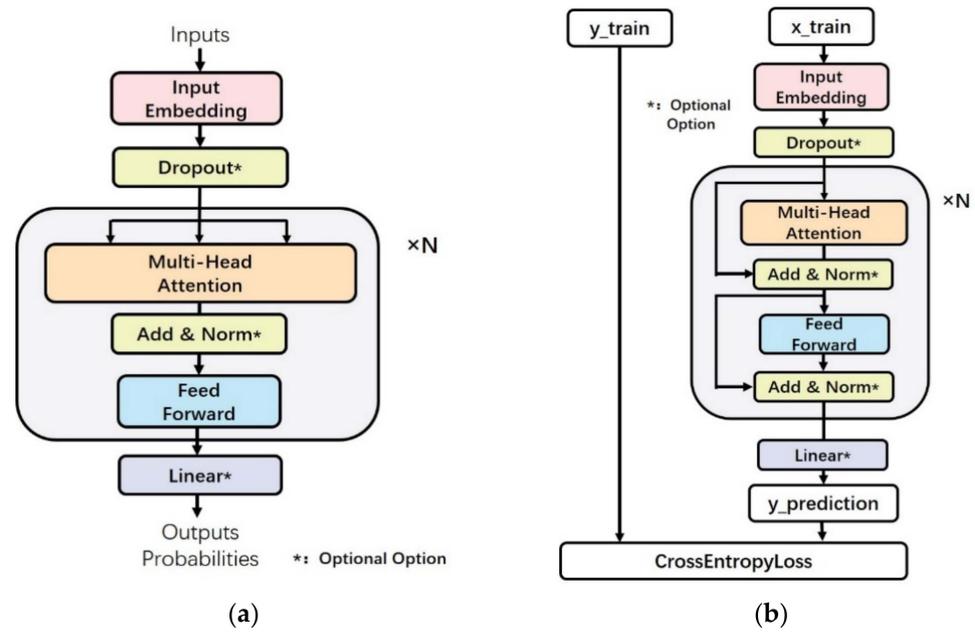
**Table 2.** Representative samples for each label of the dataset.

Longitude	Latitude	Depth	Reflection	Labels
115.0059097	20.16825028	165	0.456465	0
115.0033883	20.16775583	180	0.170332	1
114.9922325	20.16547861	251	0.107705	2
115.0038692	20.16785806	167	0.124436	3
114.9051722	20.1493725	251	0.118240	4
...	...	...	...	...

#### 4. Sediment Classification Method Based on MSC-Transformer

##### 4.1. Principle of MSC-Transformer

The architecture of MSC-Transformer consists of an Input Embedding Layer, a stack of N Transformers connected to each other, and an optional fully connected layer. Each Transformer layer consists of a multi-head self-attention layer, a feed-forward network, and an optional residual connection network. These layers are interconnected and form the architecture of the entire model. According to the characteristics of the datasets, set the input\_dim and the output\_dim. Usually, the input feature dimension of the data is the number of features in the datasets, and the output dimension is the number of types of labels. The structure of MSC-Transformer is shown in Figure 10.



**Figure 10.** (a)Structure of MSC-Transformer; (b)Training structure of MSC-Transformer.

1. Data Input

The input data is mapped from the input\_dim to the hidden\_dim through a fully connected layer. In this model, the hidden\_dim is set to 512. After that, the data enters the encoding layer of the model through an optional dropout layer that prevents overfitting, and dropout = 0.1 is set in the MSC-Transformer model.

2. Encoding layer

Each Encoder Module of the MSC-Transformer model consists of a multi-head attention layer and a feed-forward neural network. A total of 3 identical layers are stacked and connected to each other. Except for the input of the first layer and the output of the last layer, the input of each layer is the output of the previous layer, forming a chain structure connected end to end. In this model, the number of encoding layers num\_encoder\_layers = 3.

In addition, since the seabed sediment data classification problem is a tabular classification of supervised learning, the input data is the characteristics of various bottom types (such as longitude, latitude, and depth), and the output is the Label information. Therefore, since this classification is a purely encoding problem that does not require generation and situation prediction of unknown data, there is no necessity of using a Decoder. This is also confirmed in the model structure of TabTransformer.

3. Attention mechanism

The core of the MSC-Transformer model is the attention mechanism based on the Transformer model, which is a technology that can provide better model performance. Among them, the Dot-product Attention mechanism is one of the most commonly used attention mechanisms, which can more effectively save computing space and improve computing efficiency, thereby improving the training and reasoning speed of the model. At the same time, the Dot-product attention mechanism can also help the model better understand the input data and improve the accuracy and generalization ability of the model. Therefore, this model MSC-Transformer uses the Dot-product attention mechanism. The diagram of the Dot-product attention mechanism is shown in Figure 11.

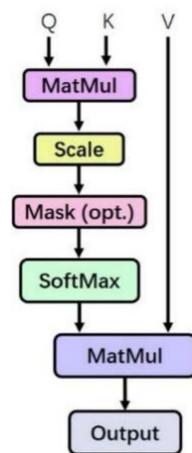


Figure 11. Mechanism of dot-production attention.

The attention function in Transformer can be described as mapping a pair of Key-Values and Query to the output and specifying the Key (K), Value (V) and Query (Q) here. The dimensions of (Q) are all  $d_k$ . In order to avoid that when  $d_k$  is large, the size of the dot product will also become large, thereby pushing the Softmax function to the region of extremely small gradients, so it is divided by the scaling factor  $\sqrt{d_k}$  to counteract this effect. The calculation formula for attention is as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

Among them, T means transpose, and Q, K, and V at this time are vectors after encoding and calculation. In the MSC-Transformer model, the number of attention heads  $\text{num\_heads} = 4$ .

#### 4. Multi-head Attention Mechanism

When the number of attention layers changes, the feature extraction ability of the data will also change accordingly, which is especially evident in the calculation and projection of weight information. In order to better mine the deep information in the classification data of the seabed sediment, the MSC-Transformer model adopts a multi-head attention mechanism. This mechanism divides the input data into multiple heads, each head calculates the attention separately, and each attention layer is distributed in parallel with each other, and finally the calculation results are stitched together through the Concat operation to achieve more refined features extract. The multi-head attention mechanism diagram is shown in Figure 12.

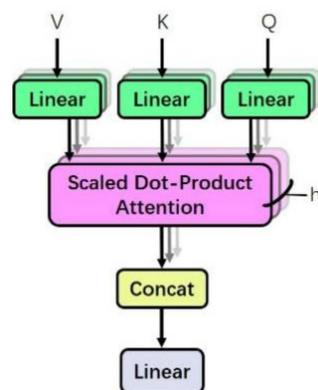


Figure 12. Mechanism of scaled dot-product attention.

The multi-head attention mechanism can decompose the input information and assign different Weights, realizing the distinction and extraction of different types of data. At the same time, through multiple attention layers distributed in parallel, the interaction and fusion between different features can be fully considered, thus improving the efficiency and accuracy of feature extraction. In addition, this method can effectively reduce the feature dimension and the amount of parameters, thereby avoiding the problems of overfitting and excessive calculation, making the training and reasoning of the model more efficient and stable. The calculation formula of the multi-head attention mechanism is as follows:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_i)W^O$$

$$\text{where head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V), i = 1, 2, 3, \dots \tag{2}$$

In the formula, Concat realizes the merging operation of the matrix, head<sub>*i*</sub> is the attention head to be merged, and the attention head follows the operation of the attention mechanism, W<sub>*i*</sub><sup>Q</sup>, W<sub>*i*</sub><sup>K</sup> and W<sub>*i*</sub><sup>V</sup> are the Query (Q), Key (K) and Value (V) vector of weight matrices.

### 5. Data Output

After the encoder layer has processed the data, it needs to be transformed into the appropriate dimensions for output. To this end, the final output of the encoder layer is passed to an optional ReLU layer, which is a hidden layer using the ReLU activation function to increase the non-linearity of the output to improve the accuracy and effectiveness of the classification. Then, this result is passed to the fully connected layer, which maps the data from hidden\_dim to output\_dim. Finally, the model prediction value results are output to complete the prediction of seabed sediment classification data information.

## 4.2. Construction of Loss Function and Optimizer

### 4.2.1. Loss Function

The loss function used in the MSC-Transformer model is the Cross-Entropy-Loss. The Cross-Entropy-Loss is a commonly used loss function, and its essence is a concept in information theory, which can be used to measure the information difference between the model prediction result and the real value. The main feature of this loss function is that it can measure the difference between the model prediction result and the real value, and use the Backward algorithm to automatically update the model parameters, thereby improving the Precision and Accuracy of the model. The model is able to predict new data more accurately [47].

Specifically, assuming that there are *n* samples and *k* categories in a dataset, the actual label of the *i*-th is *y<sub>i</sub>*, and the predicted probability distribution for the *i*-th is *p<sub>i</sub>* = (*p<sub>i1</sub>*, *p<sub>i2</sub>*, *p<sub>i3</sub>*, ... *p<sub>ik</sub>*), then the Cross-Entropy-Loss function is:

$$\text{loss} = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^k y_{ij} \log p_{ij} \tag{3}$$

*y<sub>ij</sub>* is the label of whether the *i* belongs to the *j*, and the value is 0 or 1, and *log p<sub>ij</sub>* represents the logarithm of the probability that the model predicts that the *i* belongs to the *j*. The loss function has a certain penalty mechanism during operation. When the category predicted by the model is inconsistent with the real label, the loss function will generate a larger penalty value, otherwise it will generate a smaller penalty value. This penalty mechanism is used by the backpropagation algorithm to update the parameters of the model so that the model can more accurately predict the category of each sample.

### 4.2.2. Optimizer

In the training process of deep neural networks, the most commonly used optimizer is the Stochastic Gradient Descent (SGD) optimizer [48]. The SGD optimizer and the core stochastic gradient descent algorithm is a method with a simple mathematical form but

good performance. It can effectively avoid large fluctuations in the gradient training process by adjusting the learning rate. This method is based on the assumption of independent and identical distribution. It doesn't need to calculate the gradient values of all samples every time, but updates all parameters by randomly selecting a sample. The mathematical formula for this process is as follows:

$$\begin{cases} \vec{g} = \nabla \theta L(\vec{x}_i, \vec{y}_i; \vec{\theta}) \\ \vec{\theta} \leftarrow \vec{\theta} - \eta \vec{g} \end{cases} \quad (4)$$

In this formula,  $(x_i, y_i \in D)$ ,  $\eta$  is the learning rate.

This model uses the SGD optimizer that includes momentum [49]. The momentum optimization algorithm can accelerate the convergence speed of the SGD algorithm when the gradient is updated and prevent excessive oscillation and escape from local minima. Momentum is a physical concept that expresses the product of an object's mass and velocity. In the momentum algorithm, assuming that the gradient  $g$  of the parameter is the variation of the particle velocity, the mathematical expression of the momentum algorithm can be obtained as:

$$\begin{cases} \vec{v}^t = \mu \vec{v}^{t-1} + \vec{g}^t \\ \vec{\theta}^t = \vec{\theta}^{t-1} - \eta \vec{v}^t \end{cases} \quad (5)$$

In this formula,  $\mu$  is the momentum factor.

It can be seen from the above formula that in the iterative process, the parameter update direction is not only determined by the current gradient, but also affected by the previously accumulated gradient descent direction. If the directions of the two are the same, the current gradient will be strengthened; if the directions of the two are not consistent, the gradient magnitude of the current decline will be weakened. This has the effect of accelerating convergence and reducing oscillations.

SGD itself has the nature of using only one sample to calculate the gradient at a time, and its calculation amount has nothing to do with the size of the dataset, but only with the Learning Rate, so it has a significant effect when processing a certain Batch. The advantage of it is that it can process large-scale datasets in a short period of time, reducing the amount of calculation and storage requirements. At the same time, the SGD with the momentum optimization algorithm has the effect of accelerating convergence and reducing shock, and the parameters can be updated in a larger range. This feature enables the SGD with momentum to jump out of the local optimal solution, so as to better explore the model parameter space as well as gradually optimize the model performance during the training process.

#### 4.3. The Scalability of the MSC-Transformer

The MSC-Transformer model is scalable and can maintain good training and experimental results when dealing with different numbers of seabed substrate feature classifications. In order to meet the processing of different numbers of features, it is only necessary to add the corresponding columns in the .csv file as the input data. If the total number of columns including feature columns and label columns in .csv is  $N$ , the model will automatically extract the data from the 1-st column to the  $N-1$ -th column and delete the header of the first row as the model input data; and extract the last column and delete the Header as seabed sediment classification.

In addition, the model can adjust the number of input classifications, and adjust the output\_dim provided in the model according to the number of labels of the seabed sediment classification data.

## 5. Experiment and Analysis

### 5.1. Experimental Environment and Parameter Setting

This paper uses a machine equipped with NVIDIA GeForce RTX 3090 GPU and Inter Core i5-9600 k CPU. Pytorch version is 1.13, Cuda version is 11.7, python version is 3.7.

In the process of model construction and testing, we obtained parameters and hyperparameters that make the model achieve excellent results through experiments for the dataset used in the study.

The model parameters of MSC-Transformer used in this paper are shown in Table 3.

**Table 3.** Parameters of MSC-Transformer.

Parameters	Values
batch_size	300
num_encoder_layers	3
hidden_dim	512
num_heads	4
dropout	0.1

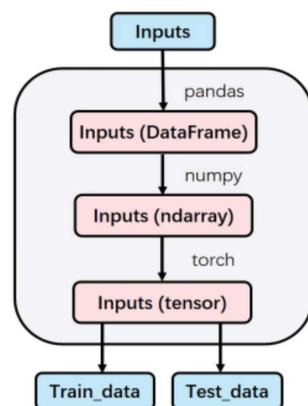
In the process of model training, the configuration of relevant hyperparameters is shown in Table 4.

**Table 4.** Hyperparameters of MSC-Transformer.

Parameters	Configuration
Epochs	3000
Learning Rate	0.0005
Momentum	0.09

### 5.2. Data Division and Model Training

The most important data structure in the deep learning framework is Tensor, which is an array that supports efficient scientific operations and can be used to store model parameters, input data, and output results [49]. The input data received by the MSC-Transformer model is in tensors. Before the data are imported into the model for calculation, it needs to undergo data structure conversion. The conversion process involves the transformation of data type, shape and dimension to meet the input requirements of the model. The conversion process is shown in Figure 13.



**Figure 13.** Conversion of inputs.

In machine learning and deep learning, in order to meet the needs of model training and evaluation, the entire dataset is usually divided into two parts: training set and test

set. The training set is used to train the model, and the test set is used to evaluate the performance of the model.

By dividing the dataset into training set and test set, the generalization ability of the model can be effectively evaluated, that is, whether the model can make accurate predictions on unseen data. This paper uses the `model_selection.train_test_split` method in the Scikit-Learning, and sets the parameters `train_size = 0.7` and `test_size = 0.3` [47].

For the training set data, train after importing the model. The training process is shown in Figure 14.

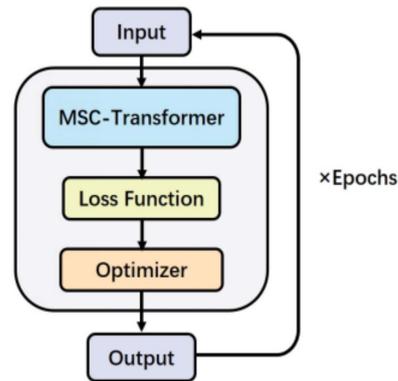


Figure 14. Model training.

For the test set data, after importing the trained model, the test process is shown in Figure 15.

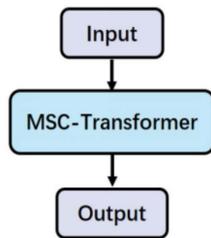


Figure 15. Model testing.

### 5.3. Experiment Results and Analysis

Figure 16 shows the statistics of loss values when MSC-Transformer uses the seabed sediment dataset in 10 independent repeated experiments.

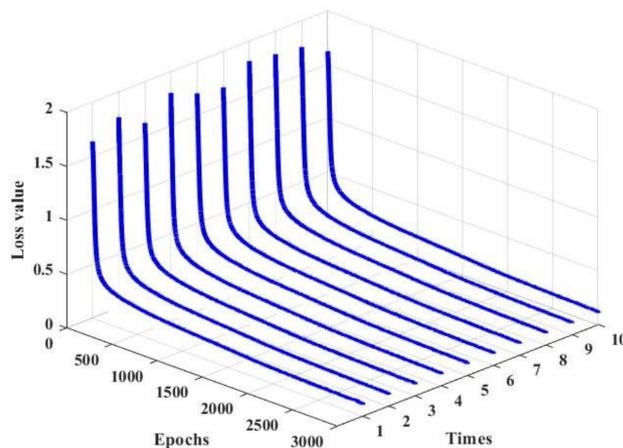


Figure 16. Loss values of 10 independent repeated experiments. The blue lines represent each variation of loss values respectively.

It can be seen from Figure 13 that before the number of training times reaches 500, the value of the loss function shows a downward trend and the rate of decline is rapid; when the number of training times is between 500 and 1500 times, the value of the loss function shows a downward trend and the decreasing. The rate has eased compared to the first 500 times; when the number of training times is greater than 1500 times, the value of the loss function decreases slowly until it converges.

This paper uses Accuracy, F1, Recall, and Precision to characterize the performance of the model. Accuracy is the ratio of the number of samples correctly classified by the model on the test set to the total number of samples. The recall is the proportion of correctly predicted positive samples to all actual positive samples. Precision is the proportion of the number of correctly predicted positive samples to the total number of predicted positive samples. F1 is the result of comprehensively considering Precision and Recall and taking their harmonic mean. Generally, the higher the values of these 4 evaluation indicators, the better the performance of the representative model, and the predicted results are closer to the real seabed sediment situation.

The Accuracy, F1, Recall, and Precision of MSC-Transformer’s 10 independent repeated experiments are shown in Figures 17–20.

Compared with BP, CNN, and RNN, MSC-Transformer conducted multiple experiments and found that the differences and fluctuations in Accuracy, F1, Recall, and Precision were small, maintaining good stability. This laid a good foundation for the accurate reproduction of the model.

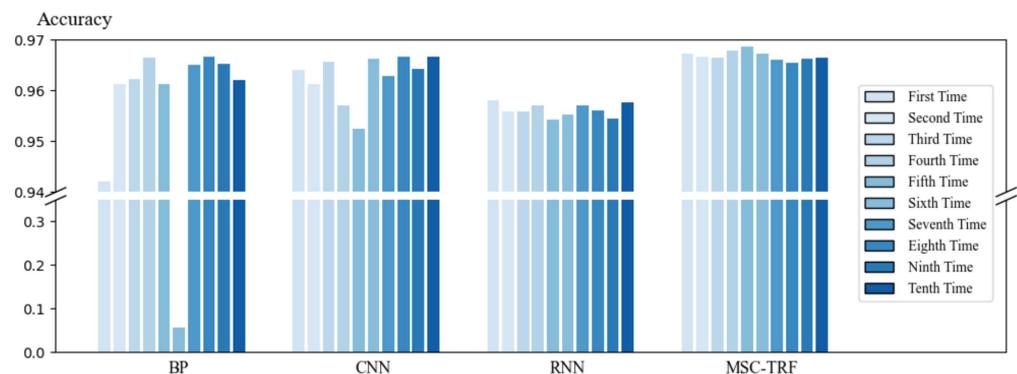


Figure 17. Accuracy of 10 independent repeated experiments.

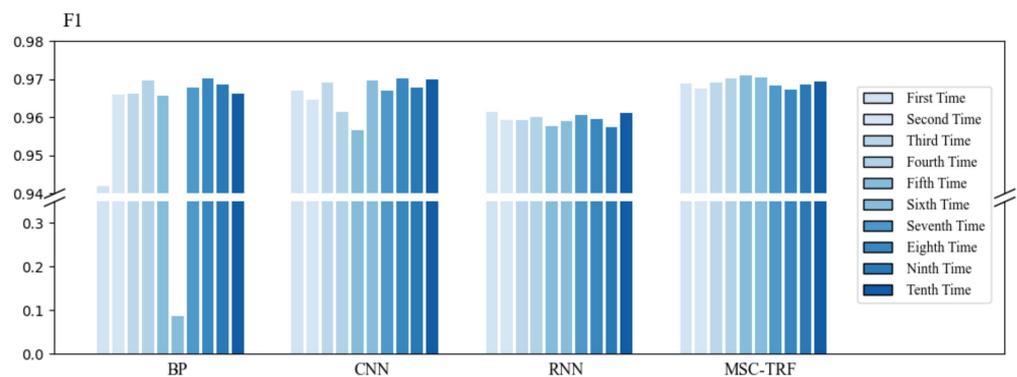


Figure 18. F1 of 10 independent repeated experiments.

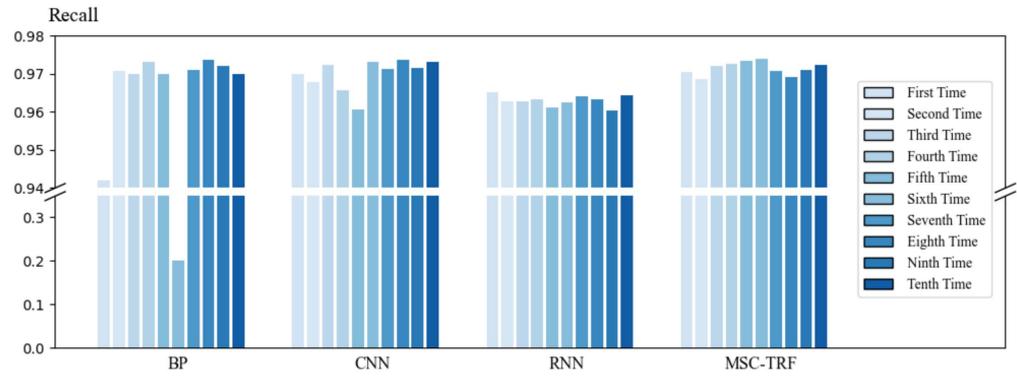


Figure 19. Recall of 10 independent repeated experiments.

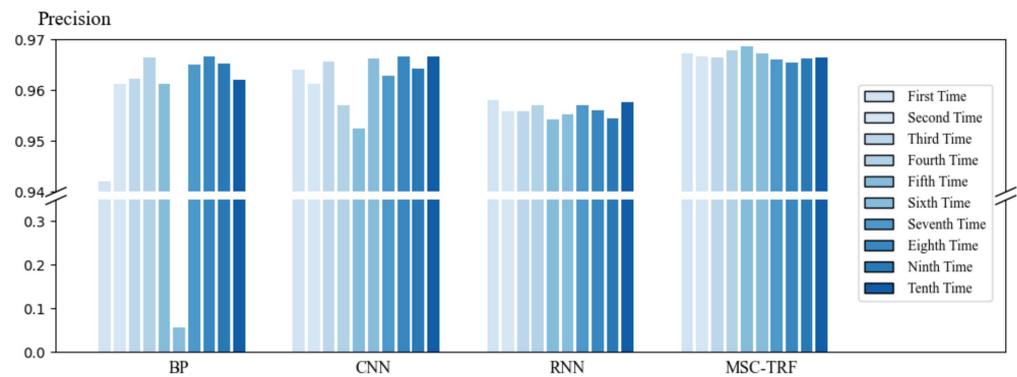


Figure 20. Precision of 10 independent repeated experiments.

In addition, Figure 21 is the result of Accuracy, F1, Recall, and Precision at the seabed sediment dataset when we compared MSC-Transformer with commonly used BP in machine learning and CNN, RNN in deep learning, and took 10 independent repeated experiments and the average value.

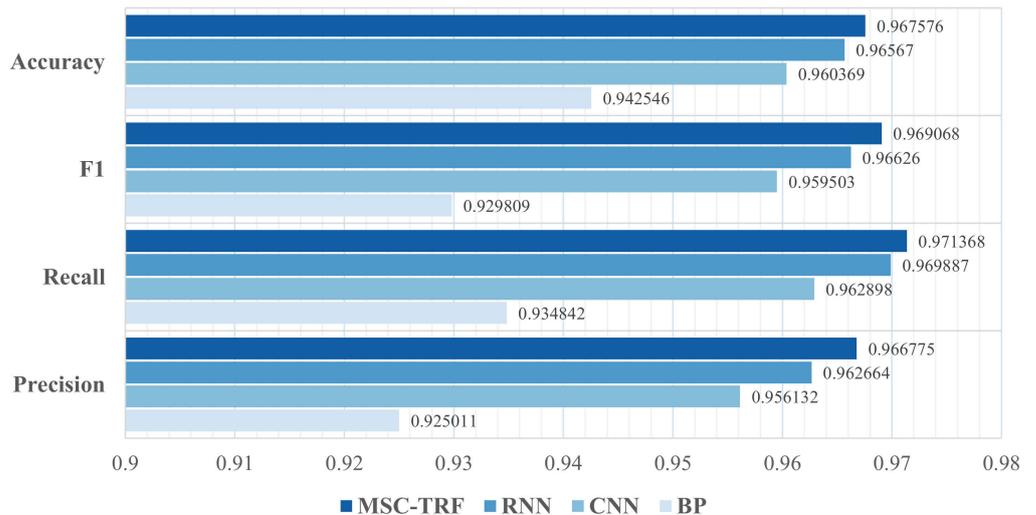


Figure 21. Accuracy, F1, Recall and Precision of Each each model.

In addition, compare MSC-Transformer with commonly used BP in machine learning and CNN, RNN in deep learning, and take 10 independent repeated experiments and the average value, Figure 21 is the result of Accuracy, F1, Recall, and Precision at the seabed sediment dataset.

It can be seen from Figure 18. In the seabed sediment dataset, MSC-Transformer has significant advantages in Accuracy, F1, Recall, and Precision compared to BP, CNN, and RNN. And the fluctuation degree of MSC-Transformer is relatively small, which reflects the stability and generalization of MSC-Transformer.

## 6. Conclusions

Seabed sediment is an important factor in the marine environment, and it has always been a major focus and difficulty in the field of marine environment research. In order to enhance and improve the classification quality of seafloor sediments, this paper applies the Transformer to marine sediment classification for the first time, proposes an MSC-Transformer based on deep learning and Transformer framework, trains and verifies the feasibility and accuracy of the model using data from the northern slope of the South China Sea obtained by a sub-bottom profiler. Through experiments, the conclusions can be drawn as follows:

- (1) The improved Transformer can be combined with data obtained from sub-bottom profilers and applied to seabed sediment classification. MSC-Transformer proposed in this paper has shown good performance in sediment classification in the northern slope of the South China Sea, with Accuracy, F1, Recall, and Precision all four evaluation indicators above 96%.
- (2) In the experiment, this paper compares MSC-Transformer with common machine learning and deep learning neural networks such as BP, CNN, and RNN. The results showed that the Accuracy, F1, Recall, and Precision of MSC-Transformer were all higher than other neural networks. Meanwhile, in 10 independent repeated experiments, MSC-Transformer showed the smallest fluctuation in all indicators and high stability, which is more conducive to seabed sediment classification in practical applications.

In addition, based on the characteristics of the model in the construction, training, and application process in this article, as well as the characteristics of the dataset of the northern slope of the South China Sea, further research directions are as follows:

- (1) Some scholars believe that one of the cores of the Transformer model is the complete Encoder-Decoder system. If you ignore the Decoder and only use the Encoder to complete the classification task of the seabed, this cannot be regarded as a formal Transformer. This paper tries to use the output parameters and input parameters of the Encoder, and add them to the Decoder to build a model. However, the model does not work properly. Therefore, under the premise of not affecting the prediction information of the seabed sediment classification, the MSC-Transformer model in this paper discards the Decoder. However, whether it is actually possible to achieve results by adding a Decoder remains to be further studied.
- (2) In the process of processing the types of seabed sediment classification, the selection of the learning rate of the model needs further study. The learning rate currently used is obtained through experiments, which can make the current model show good performance. Whereas, whether there is a more scientific method to calculate the learning rate requires further exploration by researchers in related industries.
- (3) Since the data used in the experiment are all collected from the northern slope of the South China Sea, the optimal parameters of the model are also debugged based on this dataset and good results have been achieved. Nevertheless, whether this model is applicable to datasets of other sea areas and substrate types needs to be further verified.

In summary, the MSC-Transformer in this article has excellent results under the four evaluation indicators of Accuracy, F1, Recall, and Precision, which confirms the feasibility and effectiveness of the model. The model proposed by this research institute provides a high-precision and high-tech classification tool for seabed sediment classification, providing important support for the field of marine science research. It has important practical application value for the development of marine technology, the construction of marine engineering, and the security of national defense.

**Author Contributions:** Conceptualization, X.Z. (Xiaobo Zhang), H.W. and Q.Z.; methodology, H.W., S.W., X.X. and X.Z. (Xiaobo Zhang); software, H.W.; validation, H.W., Q.Z., S.W. and X.X.; data curation, Q.Z., S.W. and X.X.; writing—original draft preparation, H.W.; writing—review and editing, H.W., S.W., X.X. and X.Z. (Xinghua Zhou); supervision, X.Z. (Xiaobo Zhang). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (No. 42106072), the Basic Scientific Fund for National Public Research Institutes of China (2021Q03), the Shandong Provincial Natural Science Foundation (No. ZR2020QD071).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Tang, Q.; Liu, B.; Chen, Y.; Zhou, X. Seabed Classification with Improved BP Neural Network. *Hydrogr. Surv. Charting* **2009**, *29*, 40–43+56. [[CrossRef](#)]
2. Kenny, A.J.; Cato, I.; Desprez, M.; Fader, G.; Schüttenhelm, R.T.E.; Side, J. An Overview of Seabed-Mapping Technologies in the Context of Marine Habitat Classification. *ICES J. Mar. Sci.* **2003**, *60*, 411–418. [[CrossRef](#)]
3. Zheng, H.; Yan, P.; Chen, J.; Wang, Y. Seabed Sediment Classification in the Northern South China Sea Using Inversion Method. *Appl. Ocean. Res.* **2013**, *39*, 131–136. [[CrossRef](#)]
4. Chen, J.; Chen, L.; Xia, X.; Zhu, R. Classification of Seafloor Sediments Sonar Image Based on Transfer Learning. *Comput. Integr. Manuf. Syst.* **2022**, *39*, 229–233. [[CrossRef](#)]
5. Blott, S.J.; Pye, K. Particle Size Scales and Classification of Sediment Types Based on Particle Size Distributions: Review and Recommended Procedures. *Sedimentology* **2012**, *59*, 2071–2096. [[CrossRef](#)]
6. Lou, Z.; Zhou, C.; Wu, C. Advances and challenges in seabed sediment classification. In *Proceedings of the 2020 Joint Academic Annual Meeting of Chinese Geosciences (26), Chongqing, China, 18 October–21 October 2020*; Beijing Botong Electronic Publishing House: Beijing, China, 2020; Volume 2. [[CrossRef](#)]
7. Zhao, M.; Hu, C. Single parameter inversion and seabed sediment classification in frequency domain. *Acta Acust.* **2021**, *46*, 1124–1131. [[CrossRef](#)]
8. Hou, Z.; Guo, C.; Wang, J.; Li, H.; Li, T. Tests of New In-Situ Seabed Acoustic Measurement System in Qingdao. *Chin. J. Ocean. Limnol.* **2014**, *32*, 1172–1178. [[CrossRef](#)]
9. Pan, G. Remote Acoustic Classification of Seafloor Sediments: A Review. *Ocean. Technol.* **1997**, *6*, 16–21.
10. Mitchell, N.C.; Clarke, J.E.H. Classification of Seafloor Geology Using Multibeam Sonar Data from the Scotian Shelf. *Mar. Geol.* **1994**, *121*, 143–160. [[CrossRef](#)]
11. Chakraborty, B.; Schenke, H.W.; Kodagali, V.; Hagen, R. Seabottom Characterization Using Multibeam Echosounder Angular Backscatter: An Application of the Composite Roughness Theory. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 2419–2422. [[CrossRef](#)]
12. Tamsett, D. Sea-Bed Characterisation and Classification from the Power Spectra of Side-Scan Sonar Data. *Mar. Geophys. Res.* **1993**, *15*, 43–64. [[CrossRef](#)]
13. Rzhannov, Y.; Fonseca, L.; Mayer, L. Construction of Seafloor Thematic Maps from Multibeam Acoustic Backscatter Angular Response Data. *Comput. Geosci.* **2012**, *41*, 181–187. [[CrossRef](#)]
14. Brown, C.J.; Blondel, P. The Application of Underwater Acoustics to Seabed Habitat Mapping. *Appl. Acoust.* **2009**, *70*, 1241. [[CrossRef](#)]
15. Preston, J. Automated Acoustic Seabed Classification of Multibeam Images of Stanton Banks. *Appl. Acoust.* **2009**, *70*, 1277–1287. [[CrossRef](#)]
16. Anderson, J.T.; Van Holliday, D.; Kloser, R.; Reid, D.G.; Simard, Y. Acoustic Seabed Classification: Current Practice and Future Directions. *ICES J. Mar. Sci.* **2008**, *65*, 1004–1011. [[CrossRef](#)]
17. Reut, Z.; Pace, N.G.; Heaton, M.J.P. Computer Classification of Sea Beds by Sonar. *Nature* **1985**, *314*, 426–428. [[CrossRef](#)]
18. Buscombe, D.; Grams, P.E.; Smith, S.M.C. Automated Riverbed Sediment Classification Using Low-Cost Sidescan Sonar. *J. Hydraul. Eng.* **2016**, *142*, 06015019. [[CrossRef](#)]
19. Atallah, L.; Probert Smith, P.J.; Bates, C.R. Wavelet Analysis of Bathymetric Sidescan Sonar Data for the Classification of Seafloor Sediments in Hopvågen Bay-Norway. *Mar. Geophys. Res.* **2002**, *23*, 431–442. [[CrossRef](#)]
20. Huseby, R.B.; Milvang, O.; Solberg, A.S.; Bjerde, K.W. Seabed Classification from Multibeam Echosounder Data Using Statistical Methods. In *Proceedings of the OCEANS '93 IEEE Conference, Victoria, BC, Canada, 18–21 October 1993*; Volume 3, pp. 229–233. [[CrossRef](#)]

21. Pearman, T.R.R.; Robert, K.; Callaway, A.; Hall, R.; Lo Iacono, C.; Huvenne, V.A.I. Improving the Predictive Capability of Benthic Species Distribution Models by Incorporating Oceanographic Data—Towards Holistic Ecological Modelling of a Submarine Canyon. *Prog. Oceanogr.* **2020**, *184*, 102338. [[CrossRef](#)]
22. Tang, Q.; Lei, N.; Li, J.; Wu, Y.; Zhou, X. Seabed Mixed Sediment Classification with Multi-Beam Echo Sounder Backscatter Data in Jiaozhou Bay. *Mar. Georesour. Geotechnol.* **2015**, *33*, 1–11. [[CrossRef](#)]
23. Tang, Q.; Ji, X.; Ding, J.; Zhou, X.; Li, J. Research prospect of acoustic seabed classification using multibeam echo sounder. *Adv. Mar. Sci.* **2019**, *37*, 1–10.
24. Li, M. Research on Sub-Bottom Profile Data Processing and Seabed Sediment Classification. Master's Thesis, Tianjin University, Tianjin, China, 2021. [[CrossRef](#)]
25. Zhang, T.; Qi, S.; Wang, X.; Tang, J. Technical Status and Development Trend of Deep Sea Sub-Bottom Profiler. *Chin. J. Eng. Geophys.* **2018**, *15*, 547–554.
26. Yang, G.; Zhu, J.; Zhao, D.; Xiong, Z.; Wang, C.; Ou, X.; Jia, Y.; Li, S. Development and application of sub-bottom profiler technologies. *Mar. Sci.* **2021**, *45*, 147–162.
27. Schock, S.G.; LeBlanc, L.R.; Mayer, L.A. Chirp Subbottom Profiler for Quantitative Sediment Analysis. *Geophysics* **1989**, *54*, 445–450. [[CrossRef](#)]
28. Gutowski, M.; Bull, J.; Henstock, T.; Dix, J.; Hogarth, P.; Leighton, T.; White, P. Chirp Sub-Bottom Profiler Source Signature Design and Field Testing. *Mar. Geophys. Res.* **2002**, *23*, 481–492. [[CrossRef](#)]
29. LeBlanc, L.R.; Mayer, L.; Rufino, M.; Schock, S.G.; King, J. Marine Sediment Classification Using the Chirp Sonar. *J. Acoust. Soc. Am.* **1992**, *91*, 107–115. [[CrossRef](#)]
30. Stevenson, I.R.; McCann, C.; Runciman, P.B. An Attenuation-Based Sediment Classification Technique Using Chirp Sub-Bottom Profiler Data and Laboratory Acoustic Analysis. *Mar. Geophys. Res.* **2002**, *23*, 277–298. [[CrossRef](#)]
31. Alevizos, E.; Snellen, M.; Simons, D.; Siemes, K.; Greinert, J. Multi-Angle Backscatter Classification and Sub-Bottom Profiling for Improved Seafloor Characterization. *Mar. Geophys. Res.* **2018**, *39*, 289–306. [[CrossRef](#)]
32. Saleh, M.; Rabah, M. Seabed Sub-Bottom Sediment Classification Using Parametric Sub-Bottom Profiler. *NRIAG J. Astron. Geophys.* **2016**, *5*, 87–95. [[CrossRef](#)]
33. Zhou, X.; Chen, Y. Seafloor Sediment Classification Based on Multibeam Sonar Data. *Geo-Spat. Inf. Sci.* **2004**, *7*, 290–296. [[CrossRef](#)]
34. Pang, Y.; Xu, F.; Liu, J. Seabed sediment classification based on gammatone filter banks time-frequency spectrum and convolutional neural networks. *Appl. Acoust.* **2021**, *40*, 510–517.
35. Gan, Y.; An, J.; Xu, X. Survey of Short Text Classification Methods Based on Deep Learning. *Comput. Eng. Appl.* **2023**, *59*, 43–53.
36. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
37. Yang, F.; Zhu, Z.; Li, J.; Feng, C.; Xing, Z.; Wu, Z. Seafloor classification based on combined multibeam bathymetry and backscatter using deep convolution neural network. *Acta Geod. Cartogr. Sin.* **2021**, *50*, 71–84.
38. Berthold, T.; Leichter, A.; Rosenhahn, B.; Berkhahn, V.; Valerius, J. Seabed Sediment Classification of Side-Scan Sonar Data Using Convolutional Neural Networks. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 27 November–1 December 2017; IEEE: Honolulu, HI, USA, 2017; pp. 1–8.
39. Luo, X.; Qin, X.; Wu, Z.; Yang, F.; Wang, M.; Shang, J. Sediment Classification of Small-Size Seabed Acoustic Images Using Convolutional Neural Networks. *IEEE Access* **2019**, *7*, 98331–98339. [[CrossRef](#)]
40. Qin, X.; Luo, X.; Wu, Z.; Shang, J. Optimizing the Sediment Classification of Small Side-Scan Sonar Images Based on Deep Learning. *IEEE Access* **2021**, *9*, 29416–29428. [[CrossRef](#)]
41. Hong, J. Summary of Transformer research status. *China CIO News* **2022**, *2022*, 125–128.
42. Wang, F. Research Progress and Application of LSTM Recurrent Neural Network. Master's Thesis, Heilongjiang University, Harbin, China, 2021.
43. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762.
44. Carion, N.; Massa, F.; Synnaeve, G.; Usunier, N.; Kirillov, A.; Zagoruyko, S. End-to-End Object Detection with Transformers. *arXiv* **2020**, arXiv:2005.12872.
45. Huang, X.; Khetan, A.; Cvitkovic, M.; Karnin, Z. TabTransformer: Tabular Data Modeling Using Contextual Embeddings. *arXiv* **2020**, arXiv:2012.06678.
46. Lin, T.; Wang, Y.; Liu, X.; Qiu, X. A Survey of Transformers. *arXiv* **2021**, arXiv:2106.04554. [[CrossRef](#)]
47. Liu, S. *Python Machine Learning Algorithm Principle, Implementation and Case*; Tsinghua University Press: Beijing, China, 2019; ISBN 978-7-302-53650-5.
48. Robbins, H.; Monro, S. A Stochastic Approximation Method. *Ann. Math. Stat.* **1951**, *22*, 400–407. [[CrossRef](#)]
49. Wang, B.; Zhou, L.; Chen, Y. *Introduction and Practice of Deep Learning Framework PyTorch*, 2nd ed.; Publishing House of Electronics Industry: Beijing, China, 2022; Volume 2, pp. 43–44. ISBN 978-7-121-43751-9.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.